DDN | Solution Brief

Accelerate >

Seismic Processing

Driving speed and precision in seismic processing with storage solutions from DDN.
Companies engaged in oil and gas exploration and production recognize the value of speed and precision. Oil and gas companies and their partners are investing tens of billions of dollars in new gas and oil development projects. With this much money at stake, there is no room for anything but the most precise and accurate information on where to drill, how much to bid on a site and how to apply the most sophisticated modeling and simulation technology to maximize reservoir performance.

This never-ending quest for speed and precision has helped to shape and define industry investments in High-Performance Computing (HPC). Competition is fierce, and companies continue to safeguard their technology as a competitive weapon, enabling them to constantly push the limits of their high-performance compute platforms. As a result, they are gathering far more data than ever and processing it at record speeds to create models of ever-increasing fidelity.

The pressure to analyze more data faster requires concurrent innovation in the underlying storage infrastructure supporting today’s seismic processing initiatives. As one of the only storage companies that has been able to successfully deliver on the capacity, speed, performance and scalability requirements for the seismic processing industry, DataDirect Networks (DDN), has contributed to major successes and breakthroughs in oil & gas exploration and production worldwide.

As an example of the capacity and performance required for today’s seismic processing solutions (discussed in greater detail below) DDN’s Storage Fusion Architecture scales up to 3.6 petabytes per rack, with performance of 1.4 million IOPS and throughput of 40 GB/second for both reads and writes. This is not your standard enterprise storage solution, but one tailored to meet the specific demands of high-performance computing environments, such as those in seismic exploration and production.

Another major differentiator for DDN is its extensive experience and expertise in file systems for massively scalable storage solutions. DDN’s solutions are open and customers can choose any of the leading HPC parallel file systems to avoid lock-in and adapt to changes in the application layer or operational environment. DDN is the leading supplier and implementer of the top HPC file systems, including Lustre, and GPFS. Over 60 percent of the world’s 100 fastest supercomputers rely on DDN storage systems, and most of them are using Lustre or GPFS.

For companies in seismic processing and exploration, the goal is to create models with ever-increasing fidelity. This enables them to be armed with the most accurate information possible, providing rapid insight on where to drill for oil or gas and how much to bid on a particular competitive site. The faster they can make the right decision, the faster they can deliver return on investment and profitability. There is a lot at stake when delays can cost millions of dollars and mistakes can cost in the tens of millions of dollars or more.
There are two fundamental ways in which oil and gas companies are pushing the boundaries of scientific discovery:

1. **Higher-resolution capture methods**, such as wide azimuth (WAZ), which uses a multi-sensor array to produce a higher-fidelity image, as well as newer approaches, such as multi-azimuth (MAZ) or rich-azimuth (RAZ). The deployment of any of these richer formats results in a vast increase in the amount of data the company must process and manage.

2. **Constant algorithmic evolution**, incorporating new analytics techniques that allow for continued advancement in the interpretation of seismic data both for new data coming in and for historical oil field data as well. This means more data to be analyzed at greater speeds.

By stretching the boundaries of scientific discovery, companies engaged in seismic exploration and processing are also pushing the limits of their underlying technology infrastructures. With the transition to these newer formats for higher-resolution capture methods, single-file sizes in excess of 100TB are not uncommon, and this data must be ingested before it can be processed and analyzed. In some cases, a dedicated seismic specialist firm will handle the data capture, and contract with an oil company. In these cases, the data still needs to be transferred, and in all cases the data needs to be stored, updated and revised.

The push for constant advance in algorithms puts even more pressure on the storage infrastructure to handle high-volume and high-velocity data. The oil and gas companies engaged in seismic processing can be much more precise in their discoveries if they can process larger data sets in shorter periods of time, so they can conduct a greater number of simulations within a specified time frame. Their models can be more accurate if these simulations can be processed and analyzed using ever-greater amounts of inputted data.

In addition, another major change in seismic processing, which is also pushing the limits of data storage, is the use of pre-stack data in routine seismic interpretation and characterization workflows. With pre-stack processing, organizations can use the entire seismic data set to create clearer understandings of earth models. This can help organizations make better business decisions in a much shorter time frame, but it puts a huge amount of pressure on the storage infrastructure. With pre-stack, “the storage requirements jump a hundred-fold and with that the stress on just about every component in the data chain: loading, storing, referencing and, most importantly, feeding to the computation/visualization node of vast, randomly searched data samples in untold amounts,” according to an article in First Break.

---

1. Data and Discovery: DDN Solutions for Seismic Processing, By Addison Snell, Intersect360 Research, September 2012
2. Data management challenges in the pre-stack era, by Philip Neri, First Break, January 2011
Seismic Processing and Big Data

The storage challenges faced by the oil & gas companies engaged in seismic exploration and production are in many ways unique to their specific requirements, but are, in other ways, somewhat representative of what other industries are experiencing in the push towards Big Data analytics. Big Data has been described by McKinsey & Company as “the next frontier for innovation, competition and productivity.” The research firm IDC has characterized Big Data as one of four pillars of computing’s next dominant platform, along with cloud computing, social networking and mobility.

Gartner has defined Big Data as a term to describe “extreme information management and processing issues that exceed the capability of traditional information technology along one or multiple dimensions to support the use of information assets.” This characterization certainly fits the challenges faced in seismic processing. In particular, the growing volume of data, the velocity of data, and the speed at which it must be processed and analyzed. In a recent survey, respondents at both HPC-oriented and non-HPC-oriented organizations cited IOPS as the top challenge for their Big Data workflows. Storage capacity ranked second in the technical rating, and I/O bandwidth and capacity were rated as the two most important storage attributes for Big Data applications.

DDN: Driving Big Data Storage Solutions for Seismic Processing

High-performance storage solutions from DDN have been designed to meet the most pressing storage challenges of the seismic processing industry for capacity, speed, performance and scalability. The key to DDN’s approach to Big Data is Storage Fusion Architecture® (SFA™), the winner of the HPCwire Editor’s Choice Award for “Best HPC Storage product or Technology in each of the past three years.

DDN | SFA12K™ combines SATA, SAS and solid-state disks into a simply managed, massively scalable, multi-petabyte platform that can be tailored to a balance of throughput and capacity. Data resides on the optimal disk technology, according to priority and access pattern, while the system sets the appropriate RAID level to maximize performance and cost-effectiveness. In the background, the system performs multiple levels of parity generation and real-time data integrity verification and error correction without impacting application performance.

As noted, DDN’s modular storage systems are built for density and consolidation. The SFA architecture supports up to 84 drives within each 4U enclosure, scaling up to 3.6 PB per rack, with a choice of InfiniBand or Fibre Channel connectivity. Leveraging the IOPS capabilities of SSDs, SFA12K is capable of 40GB/second of throughput for both reads and writes, 1.4 million disk IOPS and 1.7 million cached IOPS. The combination of modularity and performance feature allows users in seismic processing environments to configure systems that scale up and/or scale out in independent dimensions.
With this combination of I/O and capacity, this architecture provides optimal storage capabilities for a wide range of Big Data-driven applications, in particular those involved in seismic processing. For applications that demand performance with minimal latency, the SFA provides “embedded computing,” a feature that incorporates computational elements within the storage architecture itself. This computation-in-data enables the user to perform analytics and queries, or even complete algorithms, without transferring data off the disks to separate clusters of servers. By eliminating the data transfer’s latency, applications that leverage particularly large data sets, such as those in seismic processing, can see dramatic increases in performance.

Another core element of the SFA is its scale, whether it applies to single, large application servers or to multiple application servers with simultaneous access to data. The concept of performance-at-scale allows organizations involved in seismic exploration and production to consolidate systems in use by different groups, enabling greater collaboration in developing new approaches. Many organizations are already pursuing IT consolidation for its capital and operational cost efficiencies; DDN presents it as more than a cost savings: It can be a significant competitive advantage.

Moving Forward

There’s no doubt that the leaders in the oil and gas industry will continue pushing the envelope for any competitive advantage in their quest to discover the richest and most profitable reservoirs. It should be no surprise that, of all vertical markets in high-performance technical computing, it is the oil and gas companies that have the highest overall rate of internal software usage – programs and algorithms that are developed in-house. Better technology yields better information yields better results.

The shift to higher-resolution capture methods, the constant evolution of algorithms and the use of pre-stack data are drivers of breakthrough scientific discoveries, and drivers of dramatic changes in the requirements demanded from the underlying storage infrastructure. Storage solutions must be designed to deliver on the speed, precision, capacity, scalability and performance required for oil and gas companies to achieve both their scientific goals and their long-term business goals. DDN has been able to consistently deliver on all of these measures, and we are proud of our role in working with our partners and customers in seismic exploration and production to push the boundaries of discovery.

---

3 Interset360 Research HPC market advisory service,”HPC User Site Census: Primary Applications at HPTC Sites, August 2011
DDN | About Us

DataDirect Networks (DDN) is the world leader in massively scalable storage.

We are the leading provider of data storage and processing solutions and professional services that enable content-rich and high growth IT environments to achieve the highest levels of systems scalability, efficiency and simplicity. DDN enables enterprises to extract value and deliver results from their information. Our customers include the world's leading online content and social networking providers, high performance cloud and grid computing, life sciences, media production organizations and security & intelligence organizations.

Deployed in thousands of mission critical environments worldwide, DDN’s solutions have been designed, engineered and proven in the world’s most scalable data centers to ensure competitive business advantage for today's information powered enterprise.

For more information, go to www.ddn.com or call +1-800-837-2298

© 2013, DataDirect Networks, Inc. All Rights Reserved. DataDirect Networks, SFA, Storage Fusion Architecture, and SFA12K are trademarks of DataDirect Networks. All other trademarks are the property of their respective owners.